

Uses and misuses of the STROBE statement: bibliographic study

Bruno R. da Costa¹, Myriam Cevallos^{1,2}, Douglas G. Altman³, Anne W.S. Rutjes¹,
Matthias Egger¹

1. Institute of Social & Preventive Medicine (ISPM), University of Bern, Bern, Switzerland
2. CTU Bern, Bern University Hospital, Switzerland
3. Centre for Statistics in Medicine, Oxford, UK

Bruno R. da Costa, research fellow bdacosta@ispm.unibe.ch

Myriam Cevallos, research fellow mcevallos@ctu.unibe.ch

Douglas G. Altman, Professor doug.altman@csm.ox.ac.uk

Anne W.S. Rutjes, senior research fellow arutjes@ispm.unibe.ch

Matthias Egger, Professor and Head of Institute egger@ispm.unibe.ch

Manuscript word count: 2364 words

Corresponding author: Bruno R. da Costa, Institute of Social & Preventive Medicine (ISPM), Finkenhubelweg 11, CH-3012 Bern, Switzerland.

ABSTRACT

Objectives: Appropriate reporting is central to the informed application of findings from research to clinical practice. The *Strengthening the Reporting of Observational Studies in Epidemiology* (STROBE) recommendations consist of a checklist of 22 items that provide guidance on the reporting of cohort, case-control and cross-sectional studies, in order to facilitate critical appraisal and interpretation of study results. Within the framework of the revision of the STROBE recommendations, we examined the context and circumstances in which the STROBE statement was used in the past.

Design: We searched the Web of Science database in August 2010 for articles which cited STROBE and examined a random sample of 100 articles using a standardized, piloted data extraction form. The use of STROBE in observational studies and systematic reviews (including meta-analyses) was classified as appropriate or inappropriate. The use of STROBE to guide the reporting of observational studies was considered appropriate. Inappropriate uses included the use of STROBE as a tool to assess the methodological quality of studies or as a guideline on how to design and conduct studies.

Results: We identified 640 articles that cited STROBE. In the random sample of a 100 articles about half were observational studies (32%) or systematic reviews (19%). Of the 32 observational studies assessed, 26 (81%) made an appropriate use of STROBE and three uses (10%) were considered inappropriate. Among 19 systematic reviews, 10 (53%) uses were considered inappropriate.

Conclusions: The STROBE reporting recommendations are frequently used inappropriately in systematic reviews and meta-analyses as an instrument to assess the methodological quality of observational studies.

ARTICLE SUMMARY

Article Focus

- Appropriate reporting is central for the proper application of findings from clinical research into clinical practice.
- The Strengthening the Reporting of Observational Studies in Epidemiology (STROBE) recommendations aim to provide guidance to authors about how to improve the reporting of observational studies to facilitate critical appraisal and interpretation of study results.
- We examined the reasons for citing STROBE and found that most observational studies used STROBE as a reporting guideline, while about half of systematic-reviews used STROBE as a tool to assess the methodological quality of studies.

Key Messages

- Our study provides further evidence that authors of systematic reviews inappropriately use reporting guidelines to assess methodological study quality. Given the identified common misuse of STROBE, we discuss possible reasons and potential pitfalls of such misuse.

Strengths and Limitations

- We conducted a methodologically sound systematic review of the literature to address a relevant and insufficiently discussed issue concerning misuses of reporting guidelines. One of the main concerns of such misuse is the potential introduction of bias into systematic reviews and meta-analysis.
- A limitation of our findings is the fact that we only included articles which cited STROBE. This may have resulted in a selection bias since some researchers may use STROBE in their study and mention it in their manuscript but do not formally cite it.

INTRODUCTION

Appropriate reporting concerns the clear and complete presentation of what was planned, done, and found in a particular study. Inappropriate reporting makes the sound application of study findings into clinical practice difficult if not impossible. For instance, clinicians are often faced with reports of studies which do not provide enough detail of interventions for them to be implemented in clinical practice.[1] Inadequate reporting may also be a problem for the synthesis of evidence. It is common for systematic reviewers or meta-analysts to exclude otherwise eligible studies due to incomplete reporting. Given the human and financial resources needed to conduct clinical studies, it is clear that inadequate reporting has also important ethical and moral implications.[2]

To help improve the reporting of key items, reporting guidelines have been developed for various types of research design. The positive effect of reporting guidelines on the quality of reporting has been documented in several reviews.[3-7] The *Strengthening the Reporting of Observational Studies in Epidemiology* (STROBE) Statement was developed by a collaborative effort of epidemiologists, methodologists, statisticians, researchers and journal editors to improve the reporting of observational studies, including cohort, case-control and cross-sectional studies. The initiative was established in 2004 and subsequently developed in several meetings. STROBE was published in October 2007 simultaneously in several journals including leading biomedical journals as *Lancet*, *BMJ*, *Annals of Internal Medicine* and *PLoS Medicine*. The STROBE recommendations are accompanied by an explanation and elaboration document (E&E) which explains in detail the rationale of each item and provides examples of transparent reporting from published articles. To date the STROBE Statement is endorsed by over 100 journals as well as by the International Committee of Medical Journal Editors.[8]

The purpose of the STROBE Statement “is to ensure clear presentation of what was planned, done, and found in an observational study” and the authors of STROBE clearly stress that “the recommendations are not prescriptions for setting up or conducting studies, nor do they dictate methodology or mandate a uniform presentation”. [9] Notwithstanding the clear statement of the purpose of STROBE by its authors, some journal editors are concerned that the STROBE recommendations may be inappropriately used as an assessment tool to judge study quality or that

researchers may use STROBE as a guideline to set up or conduct observational studies.[10]

In the present study we examined in which circumstances and context STROBE has been cited by the researchers (how, where and why), and the background of researchers who cited STROBE. We considered the use of STROBE other than in relation to its original purpose as a reporting guideline for observational studies. Our results were presented and discussed during the August 2010 STROBE group meeting, when the group met to discuss a possible revision of the STROBE recommendations in light of the group's experience, anecdotal reports, feedback from STROBE users and new evidence.

METHODS

Literature search

On August 12th 2010, we conducted a two step literature search in the Web of Knowledge (WoK) database (www.isiknowledge.com). The first step consisted of a search for STROBE publications. Our search strategy to identify STROBE publications contained words used in the title of the STROBE article and author name (appendix 1). This first search was conducted in order to facilitate the second search, where we identified articles which had cited any of the STROBE publications using the "create citation report" tool available in the WoK database.

Article Selection

Eligibility criteria for our first search (i.e. search for STROBE publications) consisted of any STROBE publication regardless of language or version of publication (i.e. checklist or E&E). We excluded the STREGA (*Strengthening the Reporting of Genetic Association Studies*) recommendations, which is an extension of STROBE.

There were no eligibility restrictions for our second search (i.e. search of articles which cited STROBE). Any publication type, such as comments, editorials, systematic reviews, or observational research, was eligible for inclusion. We then randomly selected a sample of 100 articles from the final list of articles which cited STROBE for detailed assessment.

Data Collection

We collected the following information for STROBE publications identified through our initial search: journal in which STROBE was published, language of publication, and number of citations which STROBE received per year.

For the data extraction from articles which cited STROBE, a data extraction form was developed and piloted by two reviewers (BdC and MC) on a sample of 40 studies which were not part of the randomly selected articles. A single reviewer (BdC) then conducted data extraction on a randomly selected sample of 100 articles for the following items: verbatim text of sentence(s) which included STROBE citation; reason for citation; STROBE article which was cited; type of article which made the citation (observational research; systematic review or meta-analysis; comment, editorial, or letter to the editor; methodological article; or recommendations or narrative review); and background of researchers. Categories for “reason of citation” were defined during piloting of data extraction form and additional categories were defined during data extraction as necessary.

Classification of reasons for citation

Results of our search for articles which cited STROBE are presented narratively. Our sample of 100 randomly selected articles were classified into 8 groups according to reason for citation: STROBE used as a reporting guideline; STROBE used as a methodological (i.e. research design or conduct) guideline; STROBE used as an assessment tool of reporting quality; STROBE used as an assessment tool of methodological quality; STROBE cited to stress the importance of reporting guidelines; “unclear”; and “other”.

We considered the use of STROBE in reports of observational research to be appropriate when STROBE was used as a reporting guideline to ensure a clear and complete report of the study’s design, conduct, and findings. In this case it was clear from the text that one or more items of STROBE were used to guide the reporting of their manuscript. Our definition of appropriate use of STROBE is in line with the stated intentions of its authors, which “is solely to provide guidance on how to report research well” and “not prescriptions for designing or conducting studies”.^[9] Accordingly, we considered inappropriate the use of STROBE as a methodological guideline (i.e. guideline to design or conduct observational research or as an

assessment tool of methodological quality of publications reporting observational research). An example of what we considered inappropriate is the use of STROBE to assign an overall score of the methodological quality of a study. As stressed by the authors, STROBE was not developed to explain how research should be done, and so cannot validly be used to assess methodological quality. In box 1 we present text examples of appropriate and inappropriate uses of STROBE. The remainder of the citations which did not fall into either category were considered neutral, that is, neither appropriate nor inappropriate. Only observational research, systematic reviews or meta-analyses could be classified as appropriate or inappropriate.

For exploratory purposes, we classified articles which appropriately and inappropriately used STROBE according to the affiliation of the authors as a proxy for their background. Authors affiliated with epidemiology or public health departments were assumed to have methodological training in epidemiological research.

RESULTS

Search for STROBE publications

We identified 16 STROBE publications. Eleven of these publications were published in English, three in Spanish, and two in German. Ten of these publications were published in 2007, five in 2008, and 1 in 2009. Table 1 displays the journals where STROBE was published, number of citations received, impact factor of journal, language of publication, and date of publication.

Search for articles which cited STROBE

We identified 643 citations of STROBE since its first publication in October 2007 until August 2010. Two-hundred and four citations (32%) concerned the STROBE E&E document and 439 (68%) citations concerned the STROBE checklist. The STROBE checklist published at Lancet was the most cited, receiving 110 (17%) citations. The number of citations was associated with the impact factor of the journal as shown in table 1, the main exception being the Journal of Clinical Epidemiology, which received a large number of citations despite a lower impact factor.

Half of the articles in our sample of 100 randomly selected articles are observational studies (32%) or systematic reviews/meta-analyses (19%). Table 2 displays characteristics of these 51 research articles identified in our random sample.

Table 1. Overview of identified STROBE publications ordered by amount of citations received.

STROBE publication (Journals)	n of citations	Impact factor†	Publication language	Date of publication
Lancet	110	30.8	English	10.2007
Ann Intern Med	99	16.2	English	10.2007
Ann Intern Med (E&E)‡	97	16.2	English	10.2007
BMJ	69	13.7	English	10.2007
Plos Med (E&E)‡	62	13.1	English	10.2007
J Clin Epidemiol	51	3.0	English	04.2008
Plos Med	43	13.1	English	10.2007
Epidemiology (E&E)‡	42	5.5	English	11.2007
Epidemiology	21	5.5	English	11.2007
Prev Med	15	3.2	English	10.2007
Bull World Health Organ	14	5.3	English	11.2007
Gaceta Sanitaria	7	1.2	Spanish	03.2008
Internist	4	0.3	German	06.2008
Gaceta Sanitaria (E&E)‡	3	1.2	Spanish	03.2009
Rev Esp Salud Publica	3	-	Spanish	05.2008
Notfall & Rettungsmedizin	0	0.6	German	05.2008

† Impact factors retrieved from 2009 Journal Citation Report Science Edition

‡ E&E: Journals that published the explanation and elaboration (E&E) document that accompanies STROBE recommendations

Comments, editorials, and letters accounted for 15% of the articles, methodological articles for 8%, and recommendations and narrative reviews for 26%. Of the 32 observational studies assessed, 26 (81%) made an appropriate use of STROBE and three (10%) uses were considered inappropriate. Of the 19 systematic reviews or meta-analyses assessed, 10 (53%) made an inappropriate use of STROBE. Thus, of the 51 research articles, 26 (51%) used STROBE appropriately, and 13 (25%) used STROBE inappropriately. Figure 1 displays frequency count per article type within sample of 100 randomly selected articles.

Table 2. Characteristics of 51 classifiable articles citing STROBE.

Classification	Reason for citation	Observational research	Systematic review, Meta analysis
Appropriate	Used as a guideline for reporting	26	0
Inappropriate	Used as a guideline for Design & Conduct	3	0
	Assessment Method. Quality	0	10
Neutral	Assessment Reporting Quality	0	5
	Stressed the importance of reporting guidelines	0	1
	Other†	2	1
	Unclear	1	2
Total		32	19

†This category includes articles which cited STROBE in observational studies that used it as a guideline but it was unclear whether as a reporting or methodological guideline or articles which cited STROBE in a methodological comment.

Finally, out of the 26 studies which appropriately used STROBE, 13 (50%) had authors affiliated with epidemiology or public health departments, compared to 6 (46%) among the 13 studies with inappropriate use of STROBE.

DISCUSSION

STROBE has been cited over 600 times since its publication in 2007. Our citation analysis showed that most of the observational studies used STROBE as a guideline for reporting, whereas about half of the systematic-reviews and meta-analyses used STROBE inappropriately, as a methodological quality assessment tool.

The inappropriate interchangeable use of the terms “methodological quality” and “reporting quality” is common and may explain why some researchers used STROBE as a tool for the assessment of methodological quality or as a guideline to design and conduct observational research. Reporting quality refers to the completeness with which a study is presented and whether major items for the proper appraisal of internal and external validity of findings are clearly reported. Methodological quality refers to the appropriateness of the methods employed in the

design and conduct of epidemiological research which determines the reliability of findings (i.e. internal validity). Even though authors of STROBE clearly state that its purpose is to guide reporting of observational research, we found that it is commonly used as a methodological quality assessment tool in systematic reviews. Some researchers used STROBE as a guideline to help design and conduct observational research, which we also considered to be an inappropriate use although we acknowledge that some items of STROBE might be useful to consider when designing or conducting an observational study.

Although some items of STROBE may be related to risk of bias, many of the items are exclusively related to transparent reporting. For instance, the first item of STROBE asks researchers to indicate in the title of their manuscript the design of their study. Obviously, whether authors report the design of their study in the title of their manuscript will not influence the effect estimates of their study or compromise generalizability of their findings. To illustrate the pitfalls of using STROBE as a methodological assessment tool, we refer to the last text example we provided as an example of inappropriate use of STROBE in box 1. In this example, the authors of a systematic review transformed STROBE into a scale by assigning scores to each of its items and a final score was used to determine the eligibility of studies for inclusion into the systematic review. By doing so, authors may have introduced bias to their findings as different tools used to score methodological quality may result in different conclusions regarding the overall appropriateness of the methods employed in a particular study.[11]

There is extensive literature indicating that the misuse of findings of scientific articles is common across different fields of health care.[7, 12-21] In a similar analysis, Moher et al. reported that the Consolidated Standards of Reporting Trials (CONSORT) statement, a reporting guideline for randomized controlled trials, has also been misused by 4-10% of assessed articles as a guideline for quality assessment (7, 22). Taken together these results indicate a failure of not only authors but also of peer-reviewers in recognizing and avoiding misuse of guidelines. The peer-review process should ideally prevent misuses that are of major concern, in particular when misuses may dictate the outcomes of research as discussed above.

Box 1 - Text examples of appropriate and inappropriate use of STROBE

Based on the original purpose of STROBE, we consider the following text examples as appropriate or inappropriate use of STROBE

Appropriate

- “The reporting of this study conforms to the STROBE statement.”(23)
- “The STROBE guidelines were used to ensure the reporting of this observational study.” (24)
- “The paper is reported following the STROBE statement”(25)

Inappropriate

- “This prospective study was designed following recommendation of the STROBE statement.”(26)
- “The analysis followed the principles of the Strengthening the Reporting of Observational Studies in Epidemiology initiative.”(27)
- “The quality of selected studies was assessed using a modified version of the STROBE statement. ...The STROBE was modified by adding questions about the serologic method used to confirm dengue diagnosis, use of viral isolation, and whether the study was based on a single dengue outbreak or transmission season. Use of viral isolation increased the score whereas single outbreak studies received no additional points. The quality score was the number of items from the STROBE checklist addressed as a percentage of the total number of items applicable (minimum of 23 and maximum of 25). Studies with a quality assessment below 50% were excluded.”(28)

We observed an association between numbers of citations received and impact factor of journal in which STROBE was published. Even though the content of many STROBE publications are identical, researchers preferred to cite a STROBE article published in a journal with a higher impact factor. This association was also observed by Perneger, who concluded that citations received by an article are not merely explained by scientific merit.[29] Although reasons directing authors' choices are not clear, for some, a high impact factor may suggest the acceptance by a broader audience. The main exception to this association was the STROBE checklist published by the *Journal of Clinical Epidemiology*, which was also highly cited. This may have occurred because STROBE is particularly relevant for the daily work of the readers of this journal.

We found no clear pattern regarding affiliations of authors to an institution with a methodological focus. In particular, we found no difference in affiliations between

authors who used STROBE appropriately and authors who used STROBE inappropriately. A limitation of our findings is the fact that we only included articles which cited STROBE. This may have resulted in a selection bias since some researchers may use STROBE in their study and mention it in their manuscript but do not formally cite it. Such studies could not be identified by our search strategy. However, it is unclear to us in which direction this possible selection bias may have influenced our findings.

Citations to the STROBE checklist have been steeply increasing since its first publication. STROBE is commonly used according to its original purpose, that is, a reporting guideline of observational research. However, despite the clear statement of its objective by its authors, STROBE is misused by some researchers reporting observational studies and about half of authors reporting systematic reviews either because they used it as a tool to assess methodological quality or as a guideline on how to design and conduct observational studies. Further studies are required to define the consequences of the inappropriate use of reporting recommendations in clinical and epidemiological research, in particular the use of these checklists as an instrument to assess the methodological quality of studies.

Contributors

All authors conceptualized the ideas in the manuscript and read and approved the manuscript. BRDC developed the first draft and incorporated comments from authors for successive drafts. BRDC is the guarantor.

Competing interests

All authors are members of the group revising the STROBE statement.

Funding

BRDC is supported by a grant from ARCO foundation, Switzerland. STROBE is supported by the Swiss National Science Foundation (SNSF) and the Schweizerische Akademie der Medizinischen Wissenschaften (SAMW).

Copyright

I Bruno Roza da Costa the Corresponding Author of this article contained within the original manuscript which includes any diagrams & photographs, other illustrative material, video, film or any other material howsoever submitted by the Contributor(s) at any time and related to the Contribution ("the Contribution") have the right to grant on behalf of all authors and do grant on behalf of all authors, a licence to the BMJ Publishing Group Ltd and its licensees, to permit this Contribution (if accepted) to be published in BMJ Open and any other BMJ Group products and to exploit all subsidiary rights, as set out in the licence at: (<http://bmjopen.bmj.com//ifora/licence.pdf>)

REFERENCES

APPENDIX

Appendix 1. Search strategy for identifying STROBE articles

STEP#	SEARCH STRATEGY*
1	Title=(Strengthening the Reporting of Observational Studies in Epidemiology) AND Author=(von Elm)
2	Title=(Strengthening the Reporting of Observational Studies in Epidemiology) AND Author=(Vandenbroucke)
3	#1 OR #2
4	Title=(STREGA)
5	#3 NOT #4

*Web of Science was searched through the ISI Web of Knowledge platform (isiknowledge.com). The search was conducted on 12 August 2010.